1.3

SHORT-TERM RETROSPECTIVE LAND DATA ASSIMILATION SCHEMES

P. R. Houser *, B. A. Cosgrove, and J. K. Entin
NASA's Goddard Space Flight Center, Greenbelt, Maryland

D. Lettenmaier, G. O'Donnell
University of Washington, Seattle, Washington

K. Mitchell, C. Marshall, D. Lohmann
NOAA National Centers for Environmental Prediction, Camp Springs, Maryland

J. C. Schaake, Q. Duan
National Weather Service Office of Hydrology, Silver Spring, Maryland

E. Wood, F. Habets
Princeton University, Princeton, New Jersey

1. INTRODUCTION

Subsurface moisture and temperature and snow/ice stores exhibit persistence on various time scales that has important implications for the extended prediction of climatic and hydrologic extremes. Hence, to improve their specification of the land surface, many numerical weather prediction (NWP) centers have incorporated complex land surface schemes in their forecast models. However, because land storages are integrated states, errors in NWP forcing accumulates in these stores, which leads to incorrect surface water and energy partitioning. This has motivated the development of Land Data Assimilation Schemes (LDAS) that can be used to constrain NWP surface storages. An LDAS is an uncoupled land surface scheme that is forced primarily by observations, and is therefore less affected by NWP forcing biases. The implementation of an LDAS also provides the opportunity to correct the model's trajectory using remotely-sensed observations of soil temperature, soil moisture, and snow using data assimilation methods. The inclusion of data assimilation in LDAS will greatly increase its predictive capacity, as well as provide high-quality land surface assimilated data.

The Land Data Assimilation Schemes project is a multi-institutional research effort centered on the development of a data assimilation scheme suitable for near-real time and retrospective modeling. Through the use of land surface models (LSMs) as well as terrestrial and space-based observations, this data assimilation scheme will reduce errors in surface fluxes and storage quantities that are often present in LSM simulations (Figure 1). Most numerical weather prediction models include, and depend upon, an LSM for realistic forecasts, so LDAS will directly improve forecast accuracy. The LDAS currently operates at a 1/8th-degree resolution over the continental United States and makes use of NCEP-EDAS forcing data, NESDIS/GOES radiation data and NCEP Stage IV precipitation data. An LDAS web site--which features a real-time image generator and project information--has been created and is located at http://ldas.gsfc.nasa.gov. It should also be noted that several other papers presented during the 15th conference on Hydrology specifically describe other aspects of this LDAS research. These include 1.1, 1.2, 1.4, 1.6, 1.22 in the 15th conference on Hydrology, and P1.1 in the 16th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology.

To address a variety of land-surface research questions, LDAS is being implemented in real-time, short-term retrospective (1996-current), and long-term retrospective (50 year) modes. The short-term retrospective LDAS simulations use the same 1/8th degree LDAS grid, parameters, and atmospheric forcing approach as is used in the real-time LDAS simulations. This consistency between the short-term and real-time LDAS simulations facilitates: 1) land surface storage spin-up and drift analysis; 2) soil moisture, temperature, and snow initialization studies; 3) evaluation of new land surface theory and forcing; and 4) development and validation of land surface data assimilation techniques. An overview of the LDAS short-term retrospective forcing data, simulations are presented.

2. SHORT-TERM RETROSPECTIVE LDAS

2.1 LAND SURFACE MODELING

Recent advances in understanding soil-water dynamics, plant physiology, micrometeorology, and the hydrology that control biosphere-atmosphere interactions have spurred the development of Land Surface Models (LSMs), whose aim is to simply yet realistically represent the transfer of mass, energy, and momentum between a vegetated surface and the atmosphere (Dickinson *et al.*, 1993; Sellers *et al.*, 1986). LSM predictions are regular in time and space, but these predictions are influenced by model structure, errors in input variables and model parameters, and inadequate treatment of sub-grid scale spatial variability. Consequently, LDAS predictions will be

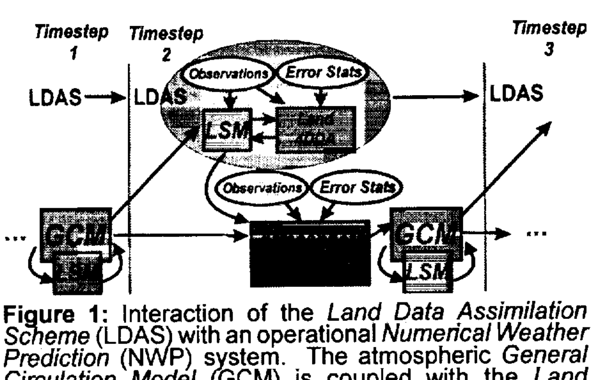

Figure 1: Interaction of the *Land Data Assimilation Scheme* (LDAS) with an operational *Numerical Weather Prediction* (NWP) system. The atmospheric *General Circulation Model* (GCM) is coupled with the *Land Surface Model* (LSM), and both use a *4-Dimensional Data Assimilation* (4DDA) process to integrate past forecasts with observations to improve performance.

* *Corresponding author address:* Paul Houser, Goddard Space Flight Center, Greenbelt, MD 20771; Paul.Houser@gsfc.nasa.gov.

influenced by the particular LSM chosen; to avoid this, several LSMs will be used in the LDAS context. These are the **Mosaic** LSM of Koster and Suarez (1992) and Koster *et al.* (1998), the **N**ational Centers for Environmental Prediction (NCEP), **O**regon State University (OSU), United States **A**ir Force (USAF), and Office of **H**ydrology (OH), LSM, called **NOAH**, and the Variable Infiltration Capacity (**VIC**) LSM of Wood *et al.* (1992).

The Mosaic LSM addresses the problem of subgrid heterogeneity by subdividing each GCM grid cell into a user-specified mosaic of tiles (after Avissar and Pielke, 1989), each tile having its own vegetation type and hence water and energy balance. Surface flux calculations for each tile are similar to those described by Sellers *et al.* (1986). Tiles do not directly interact with each other, but influence each other indirectly, by their collective influence on the overlying atmosphere. Like the plethora of LSMs that have been developed over the past decade (e.g. the PILPS participants, Henderson-Sellers *et al.* (1993)), Mosaic is well suited to modeling the vertical exchange of mass, energy and momentum with the overlying atmosphere.

The NOAA-NOAH LSM simulates soil moisture (both liquid and frozen), soil temperature, skin temperature, snowpack water equivalent, snowpack density, canopy water content, and the traditional energy flux and water flux terms of the surface energy balance and surface water balance. This model has been used in a) the NCEP-OH submission to the PILPS-2d tests for the Valdai, Russia site, b) the emerging, realtime, U.S.-domain, Land Data Assimilation System (LDAS), c) the coupled NCEP mesoscale Eta model (Chen *et al*, 1997) and the Eta model's companion 4-D Data Assimilation System (EDAS), as well as in d) the coupled NCEP global Medium-Range Forecast model (MRF) and its companion 4-D Global Data Assimilation System (GDAS).

The Variable Infiltration Capacity (VIC) model of Wood *et al.*, (1992) represents the variation in infiltration capacity within a grid cell, a drainage term for interstorm runoff, and a nonlinear evapotranspiration term. This model assumes that infiltration capacities, and therefore runoff generation and evaporation, vary within an area due to variations in topography, soil, and vegetation. The VIC model has been well tested in various model intercomparisons, GCM simulations, and sensitivity studies.

2.2 Justification for using an uncoupled LSM

There are strong justifications for studying an LSM uncoupled from atmospheric and ocean models. Coupling the LSM to an atmospheric model allows for the study of the interaction and feedbacks between the atmosphere and land surface. However, coupled modeling also imposes strong land surface forcing biases predicted by the atmospheric model on the LSM. These biases in precipitation and radiation can overwhelm the behavior of LSM physics. In fact, several NWP centers must 'correctively nudge' their LSM soil moisture toward climatological values to eliminate its drift. By using an uncoupled LSM, we can better specify land surface forcing using observations, use less computational resources, and address virtually all of the relevant scientific questions. The physical understanding and modeling insights gained from implementing distributed, uncoupled land-surface schemes with observation-based forcing has been vividly demonstrated in recent GEWEX retrospective off-line land surface modeling projects known as PILPS-2c and the Global Soil Wetness Project (Koster and Milly, 1997).

2.2 Forcing Data

The use of high quality atmospheric forcing of the land surface is essential to produce reasonable land surface predictions. This is the most daunting and important part of creating the LDAS. Land surface models require wind speed, air temperature, humidity, precipitation, and radiation on a hourly basis. Many of these land surface forcing variables can be reliably provided by operational Numerical Weather Prediction (NWP) models (i.e. the NCEP Eta model fields archived at NCAR) run in either a real-time or reanalysis mode. However, precipitation and radiation are generally poorly predicted by NWP models because we have not mastered the complex prediction of cloud physics and dynamics, which can lead to gross errors in land surface simulations. Therefore, we replace these fields by some recently emerging observational products, when available, as described below:

Surface meteorology: Gridded fields of precipitation, radiation, wind speed, air temperature are available by NOAA's National Centers for Environmental Prediction (Kalnay *et al.*, 1996) from 1996 to the present. These have been interpolated and corrected to account for the topographic difference between the 32km Eta model grid and the 1/8 degree LDAS grid.

Precipitation: The NCEP stage IV hourly 4-km gage/radar national precipitation analysis from NOAA's National Centers for Environmental Prediction form the backbone of LDAS precipitation forcing. These have been bias corrected to match 24-hour gage totals.

Radiation: NESDIS hourly 0.5 degree GOES-based surface solar insolation analysis. These have been interpolated to the LDAS grid and time step accounting for the solar Zenith angle.

Model parameters were derived from the existing high-resolution vegetation and soil coverages available from the EROS data center (see Figure 2).

3. Preliminary Results

A preliminary short-term retrospective LDAS simulation was conducted starting April 15, 1999, and ending July 31, 1999 (Figures 3 and 4). These Mosaic simulations were initialized from using Eta model land surface conditions on April 15, 1999. In the near future, the retrospective LDAS simulations will begin in 1996 and extend to near-real time. Figures 3 and 4 show the Mosaic average root zone soil moisture, deep soil temperature, sensible heat flux, and latent heat flux as an average for the month of July 1999, and time series for the Southern Great Plains (SGP) area (Oklahoma and Kansas), the desert southwestern U.S., and the mid-Atlantic states (Maryland and Virginia) are shown for the temporal extent of this simulation. Soil moisture shows realistic variability across the U.S., with a general drying trend for the SGP and Mid-Atlantic regions. The deep soil temperature systematically increases during this period. This drying and heating is clearly due to both model spin-up, and summer onset. Longer-term retrospective spin-up studies are clearly required to establish the proper equilibrium for the LDAS system. The predicted latent and sensible heat fluxes also show realistic spatial variability, with significant temporal variability in response to precipitation, as would be expected. It should be noted that the absence of GOES-derived solar insolation north of 50 degrees latitude results in a linear discontinuity. This feature helps to highlight the importance of observations for proper LDAS predictions.

4. Summary

The characterization of the spatial and temporal variability of water and energy cycles are critical to improve our understanding of land surface-atmosphere interaction and the impact of land surface processes on climate extremes. Because the accurate knowledge of these processes and their variability is important for climate predictions, most Numerical Weather Prediction (NWP) centers have incorporated land surface schemes in their models. However, errors in the NWP forcing accumulate in the surface and energy stores, leading to incorrect surface water and energy partitioning and related processes. This has motivated the NWP to impose ad hoc corrections to the land surface states to prevent this drift. A methodology under development here is to implement Land Data Assimilation schemes (LDAS), which are uncoupled models forced with observations, and are therefore not affected by NWP forcing biases. This research is being implemented in near real time using existing Land Surface Models (LSMs) by NCEP, NASA, Princeton University, and the University of Washington at a 1/8 degree (about 10 kilometer) resolution across the United States to evaluate these critical science questions. The LDAS is forced with real time output from numerical prediction models, satellite data, and radar precipitation measurements.

The short-term retrospective LDAS simulations are specifically investigating: 1) land surface storage spin-up and drift analysis; 2) soil moisture, temperature, and snow initialization studies; 3) evaluation of new land surface theory and forcing; and 4) development and validation of land surface data assimilation techniques. A real-time LDAS system is also currently in place (see http://ldas.gsfc.nasa.gov), that uses near-real time NCEP Eta model analysis fields, along with observed precipitation and radiation fields to force several different land surface models in an uncoupled mode.

5. References

Avissar, R. and R. Pielke, A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology, *Mon. Weather Rev.*, 117, 2113-2136, 1989.

Chen, Janjic, Mitchell, 1997, Boundary-Layer Meteorol, 85, 391-421.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy, 1993: Biosphere-Atmosphere Transfer Scheme (BATS) Version 1e as coupled to the NCAR Community Climate Model. NCAR Tech. Note., NCAR/TN-387+STR.

Henderson-Sellers, A., Z.-L. Yang and R. E. Dickinson, The Project for Intercomparison of Land-surface Parameterization Schemes, *Bull. Am. Met. Soc.*, 74, 1335-1349, 1993.

Kalnay, E. *et al.*, 1993: The NCEP/NCAR Reanalysis Project. NMC Office Note 401.

Koster, R. D., M. J. Suarez, A. Ducharne, P. Kumar, and M. Stieglitz, 1998, A catchment-based approach to modeling land

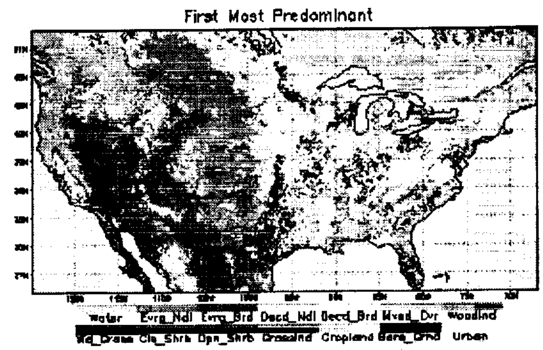

Figure 2: LDAS most predominant vegetation.

surface processes in a GCM, in review, *J. Geophys. Res.*

Koster, R. D., and P. C. D. Milly, 1997: The Interplay Between Transpiration and Runoff Formulations in Land Surface Schemes Used with Atmospheric Models. *Journal of Climate*, **10**:1578-1591.

Koster, R. D. and M. J. Suarez, 1992, Modeling the land surface boundary in climate models as a composite of independent vegetation stands, *J. Geophys. Res.*, 97, 2697-2715.

Sellers, P. J., Y. Mintz, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *Journal of Atmospheric Science*, **43**:505-531.

Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian: A Land-Surface Hydrology Parameterization With Subgrid Variability for General Circulation Models. *Journal of Geophysical Research*, 97(D3): 2717-2728.

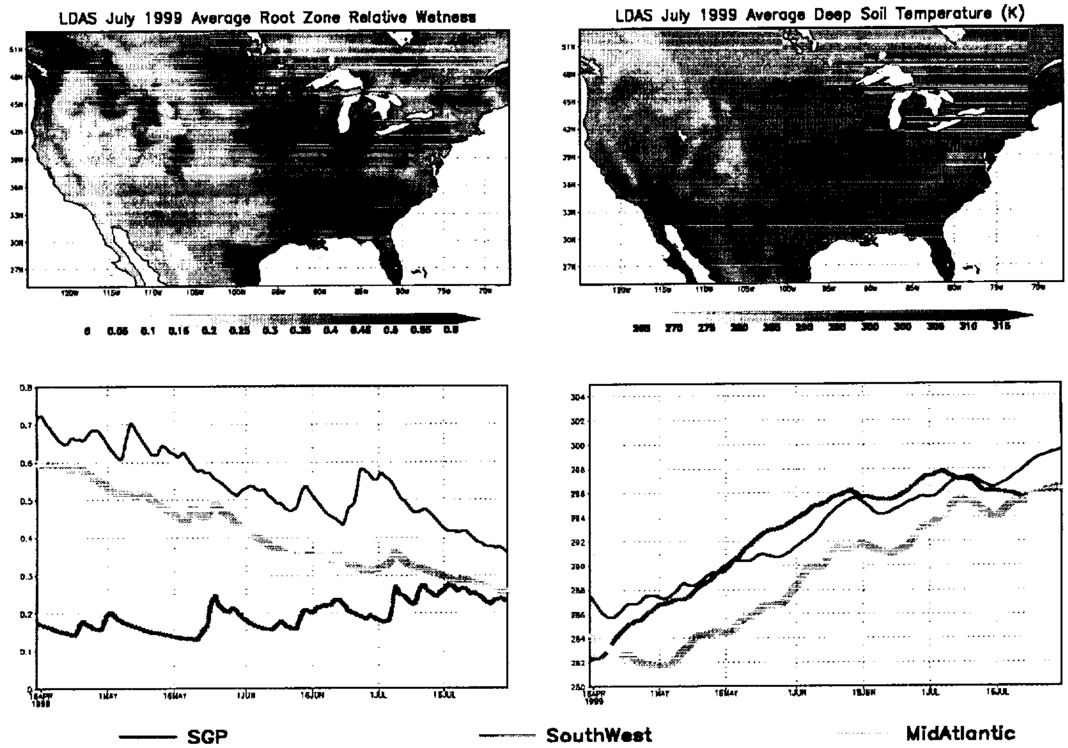

Figure 3: Upper: July average LDAS root-zone soil moisture and deep soil temperature from the Mosaic LSM. Lower: Time series of LDAS root-zone soil moisture and deep soil temperature from April 15, 1999 to July 31, 1999.

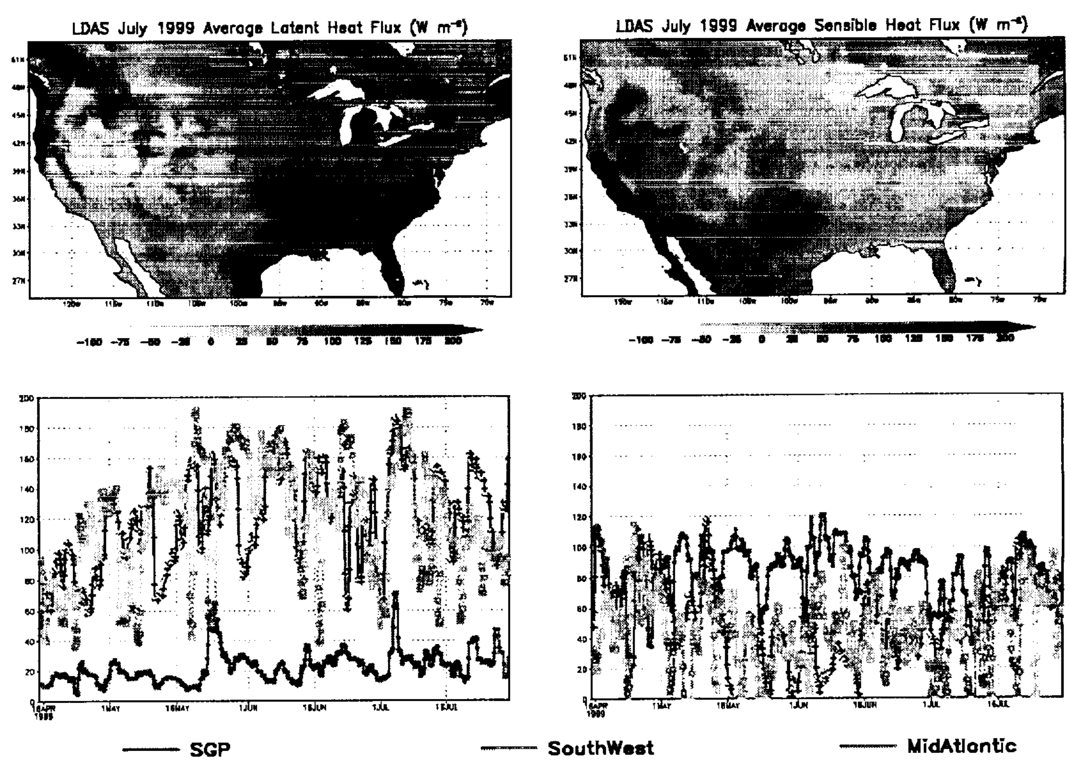

Figure 4: Upper: July average LDAS sensible and latent heat fluxes from the Mosaic LSM. Lower: Time series of sensible and latent heat fluxes for April 15, 1999 to July 31, 1999.